

Reference 1: JP 2003-157262 A

[0004]

[Means to Solve the Problems] In order to solve the above problems, a patent search apparatus comprises a patent search database section for storing a word or a compound word included in a patent document to be retrieved by associating the word or the compound word in advance with a said patent document

as a search word/phrase, a word/phrase extraction section for analyzing a search source patent document and for extracting a word or a compound word included in a patent document for setting a said condition as a search condition word/phrase, and a search section for comparing the search word/phrase and the search condition word/phrase and for retrieving a patent document highly similar to the search source patent document as a patent document nearly related to the search source patent document. With the patent search apparatus having the above structure, the word/phrase extraction section analyzes the search source patent document and extracts a word or a compound word included in the patent document for setting the said condition as a search condition word/phrase. The search section refers to the patent search database section, compares the search word/phrase and the search condition word/phrase, and retrieves a patent document highly similar to the search source patent document as a patent document nearly related to the search source patent document.

[0046] In addition, the two-dimensional distance between each of the U.S. patent publication S, the patent publication C, the patent publication B, the patent publication D, and the university paper R and the patent publication X indicates the degree of similarity. That is to say, a short distance means a high degree of similarity. A figure indicated within a rectangle indicative of each patent document means the number of patent documents which refers to a said patent document. In this example, patent publication A and patent publication E are referred to many times. Accordingly, it is possible to judge that the possibility of inventions disclosed in the patent publication A and the patent publication E being basic patents is strong. That is to say, if the patent publication X is a laid-open patent publication, an applicant for the patent will have to refer first to the inventions disclosed in the patent publication A and the patent publication E. In addition, patent publications N, Q, and P which are similar to the patent publication X exist. This means that attention must be paid to inventions disclosed in these patent publications. In this case, any patent document is selected and a search type is designated, on a display screen of the display section 103. By doing so, a search process is performed in the same way with the selected patent document as a search source patent document. By performing such a search process, a search can be continued hierarchically and desired patent documents can be retrieved efficiently.

[0047] Relationships among patent documents can be grasped more easily

by classifying line segments which connect the patent documents by color and systematically representing the patent documents, by changing color in a rectangle indicative of each patent document according to applicant, inventor, or agent (patent attorney, for example), or by using blinking for displaying a patent document. On the other hand, if a claim flag is set in step S12 (claim flag in step S12), then the claim search section 132 of the actual search processing section 114 retrieves dependence relationship between claims from the component database and the dependence relationship database (step S17). When the claim search process terminates, the chart compilation section 113 performs a process for charting claim search results (step S15). When the chart compilation section 113 completes the process for charting the claim search results, the display section 103 displays a chart corresponding to the claim search results (step S16).

FIG. 11

米国特許 S	U.S. PATENT S
特許 A	PATENT A
特許	PATENT
特許	PATENT
特許	PATENT
特許 C	PATENT C
特許	PATENT
特許 E	PATENT E
特許	PATENT
特許	PATENT
参照されている数を表示する      DISPLAY NUMBER OF PATENT	
DOCUMENTS WHICH REFER TO SAID PATENT DOCUMENT	
特許 B	PATENT B
特許 D	PATENT D
特許	PATENT
特許	PATENT
大学論文 R	UNIVERSITY PAPER R
特許 X	PATENT X
特許 N	PATENT N
特許 Q	PATENT Q
特許 P	PATENT P
特許 F	PATENT F
特許 G	PATENT G
特許	PATENT
特許	PATENT
特許	PATENT

特許 M            PATENT M

特許            PATENT

特許            PATENT

特許 I           PATENT I

特許            PATENT

特許文献 X 出願前に発表された文書      DOCUMENTS PUBLISHED  
BEFORE APPLICATION FOR PATENT ABOUT INVENTION DISCLOSED  
IN PATENT DOCUMENT X

特許文献 X 出願後に発表された文書      DOCUMENTS PUBLISHED AFTER  
APPLICATION FOR PATENT ABOUT INVENTION DISCLOSED IN  
PATENT DOCUMENT X

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-157262

(43)Date of publication of application : 30.05.2003

(51)Int.Cl.

G06F 17/30

(21)Application number : 2001-355274

(71)Applicant : SEIKO EPSON CORP

(22)Date of filing : 20.11.2001

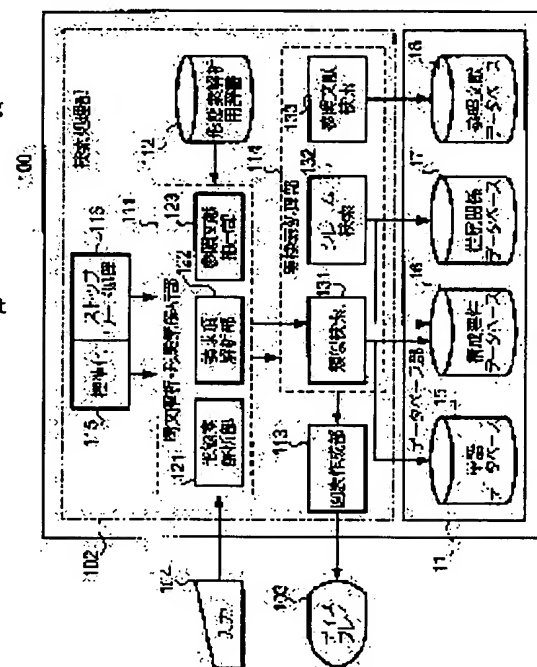
(72)Inventor : TANAKA TAKASHIGE  
YAMADA KOJI

(54) PATENT RETRIEVAL DEVICE, CONTROL METHOD THEREFOR, CONTROL PROGRAM AND RECORDING MEDIUM

(57)Abstract:

**PROBLEM TO BE SOLVED:** To clearly present referring and to-be-referred relation among a plurality of patent documents and to extract the patent document of approximate technical contents though there is not the referring relation.

**SOLUTION:** The retrieval processing part 102 of a patent retrieval system 100 analyzes a retrieving source patent document and extracts a word or a composite word included in the patent document for setting conditions as a retrieval condition word/phrase. Then, by referring to a database part 11, a word/phrase for retrieval and retrieval condition word/phrase are compared with each other to retrieve a patent document similar to the patent document of the retrieval source as a patent document highly related with the patent document of the retrieval source.



## LEGAL STATUS

[Date of request for examination]

28.09.2004

[Date of sending the examiner's decision of rejection]

12.06.2007

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

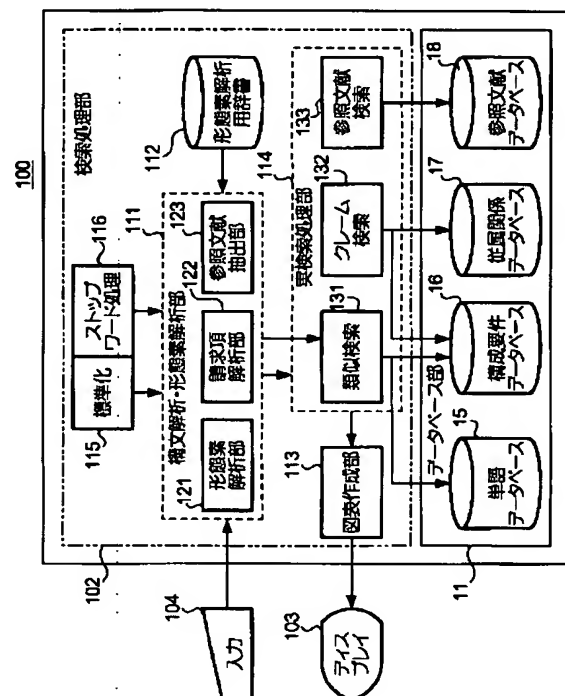
[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]



## 【特許請求の範囲】

【請求項 1】 検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部と、検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出部と、前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索部と、を備えたことを特徴とする特許検索装置。

【請求項 2】 請求項 1 記載の特許検索装置において、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出部を備えたことを特徴とする特許検索装置。

【請求項 3】 請求項 2 記載の特許検索装置において、前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備えたことを特徴とする特許検索装置。

【請求項 4】 請求項 3 記載の特許検索装置において、前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成部と、前記図表作成部により作成された図表を提示するディスプレイ部と、を備えたことを特徴とする特許検索装置。

【請求項 5】 請求項 1 ないし請求項 4 のいずれかに記載の特許検索装置において、前記検索元特許文献の形態素解析を行って前記単語あるいは前記複合語を抽出する形態素解析部と、抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度算出部と、抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前記検索用語句として登録する登録部と、を備えたことを特徴とする特許検索装置。

【請求項 6】 請求項 1 ないし請求項 5 のいずれかに記載の特許検索装置において、前記検索部は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索することを特徴とする特許検索装置。

【請求項 7】 請求項 1 ないし請求項 6 記載の特許検索装置において、前記特許文献は、複数の「請求項」の記載を含み、前記検索部は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析部を備えたことを特徴とする特許検索装置。

【請求項 8】 請求項 7 記載の特許検索装置において、前記請求項解析部は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出することを特徴とする特許検索装置。

【請求項 9】 請求項 8 記載の特許検索装置において、前記請求項解析部は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別することを特徴とする特許検索装置。

【請求項 10】 請求項 1 ないし請求項 6 のいずれかに記載の特許検索装置において、前記特許文献は、「請求項」の記載を含み、前記検索部は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析部を備えたことを特徴とする特許検索装置。

【請求項 11】 検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有する特許検索装置の制御方法において、検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出過程と、前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索過程と、を備えたことを特徴とする特許検索装置の制御方法。

【請求項 12】 請求項 11 記載の特許検索装置の制御方法において、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項 13】 請求項 12 記載の特許検索装置の制御方法において、前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備え、前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成過程と、前記図表作成部により作成された図表を提示するディスプレイ過程と、を備えたことを特徴とする特許検索装置の制御方法。

【請求項 14】 請求項 11 ないし請求項 13 のいずれかに記載の特許検索装置の制御方法において、前記検索元特許文献の形態素解析を行って前記単語ある

いは前記複合語を抽出する形態素解析過程と、抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度計算過程と、抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前記検索用語句として登録する登録過程と、を備えたことを特徴とする特許検索装置の制御方法。

【請求項15】 請求項11ないし請求項14のいずれかに記載の特許検索装置の制御方法において、前記検索過程は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索することを特徴とする特許検索装置の制御方法。

【請求項16】 請求項11ないし請求項15記載の特許検索装置の制御方法において、前記特許文献は、複数の「請求項」の記載を含み、前記検索過程は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項17】 請求項16記載の特許検索装置の制御方法において、前記請求項解析過程は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出することを特徴とする特許検索装置の制御方法。

【請求項18】 請求項17記載の特許検索装置の制御方法において、前記請求項解析過程は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別することを特徴とする特許検索装置の制御方法。

【請求項19】 請求項11ないし請求項15のいずれかに記載の特許検索装置の制御方法において、前記特許文献は、「請求項」の記載を含み、前記検索過程は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項20】 検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有するコンピュータを特許検索装置として機能させる制御プログラムであって、検索元特許文献を解析させ、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出させ、

前記検索用語句および前記検索条件語句を比較させ、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索させ

る、ことを特徴とする制御プログラム。

【請求項21】 請求項20記載の制御プログラムにおいて、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出させる、

ことを特徴とする制御プログラム。

【請求項22】 請求項20または請求項21記載の制御プログラムにおいて、

前記特許文献は、複数の「請求項」の記載を含み、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出させる、ことを特徴とする制御プログラム。

【請求項23】 請求項20または請求項21記載の制御プログラムにおいて、

前記特許文献は、「請求項」の記載を含み、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出させる、ことを特徴とする制御プログラム。

【請求項24】 請求項20ないし請求項23のいずれかに記載の制御プログラムを記録したことを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、特許検索装置、特許検索装置の制御方法、制御プログラムおよび記録媒体に係り、特に複数の特許文献相互間の関係あるいは特許文献における請求項相互間の関係を検索者に対して明確に提示するための技術に関する。

【0002】

【従来の技術】近年の技術の発達に伴い、特許出願の件数は膨大なものとなってきている。これに伴って特許文献の量も膨大となってきており、重複研究を防止し、権利侵害のチェック、出願前の調査、他社の技術開発動向の把握、研究開発の方向性の決定等を行う場合の特許文献調査も容易ではないのが実状である。このような特許文献調査における負荷を軽減するためのものとして、従来より、特許検索システムが知られている。従来の特許検索システムにおいては、検索対象の特許文献に含まれると思われるキーワードを一または複数入力し、当該キーワードあるいはキーワード群を含む特許文献をデータベースを参照して抽出し、抽出した特許文献のリストを生成し、表示あるいは出力するように構成されていた。

【0003】

【発明が解決しようとする課題】上記従来の特許検索システムにおいては、「漏れの少ない」調査を行うことができるものの、キーワードの選択を誤れば、逆に不必要な情報も膨大に含まれてしまうと言う問題点があった。そしてこのような場合には調査そのものの時間も膨大に

かかると言う問題点があった。また、不必要にキーワードを増加して絞り込みを行ってしまうと関連する特許文献も抜け落ちてしまうという問題点があった。また、複数の特許文献が抽出される場合であっても、各特許文献相互の関係は、文献内容を詳細に読み取らなければ分からないと言う問題点があった。すなわち、複数の関連する特許文献が抽出されてもどの特許文献がより基本特許に近いものであるかは、当該複数の文献に記載されている情報がある程度参照する必要がある。そこで、本発明の目的は、複数の特許文献間の参照、被参照関係を明確に提示し、さらに参照関係がなくても近似する技術内容の特許文献を抽出することが可能な特許検索装置、特許検索装置の制御方法、制御プログラムおよび記録媒体を提供することにある。

【0004】

【課題を解決するための手段】上記課題を解決するため、特許検索装置は、検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部と、検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出部と、前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索部と、を備えたことを特徴としている。上記構成によれば、特許検索装置の語句抽出部は、検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する。検索部は、特許検索データベース部を参照して前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する。

【0005】この場合において、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出部を備えるようにしてもよい。また、前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備えるようにしてもよい。さらに前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成部と、前記図表作成部により作成された図表を提示するディスプレイ部と、を備えるようにしてもよい。

【0006】さらにまた、前記検索元特許文献の形態素解析を行って前記単語あるいは前記複合語を抽出する形態素解析部と、抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度算出部と、抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前

記検索用語句として登録する登録部と、を備えるようにしてもよい。また、前記検索部は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索するようにしてもよい。

【0007】さらに前記特許文献は、複数の「請求項」の記載を含み、前記検索部は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析部を備えるようにしてもよい。さらにまた、前記請求項解析部は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出するようにしてもよい。また、前記請求項解析部は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別するようにしてもよい。さらに前記特許文献は、「請求項」の記載を含み、前記検索部は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析部を備えるようにしてもよい。

【0008】また、検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有する特許検索装置の制御方法において、検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出過程と、前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索過程と、を備えたことを特徴としている。この場合において、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出過程を備えるようにしてもよい。

【0009】また、前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備え、前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成過程と、前記図表作成部により作成された図表を提示するディスプレイ過程と、を備えるようにしてもよい。さらに前記検索元特許文献の形態素解析を行って前記単語あるいは前記複合語を抽出する形態素解析過程と、抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度計算過程と、抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前記検索用語句として登録する登録過程と、を備えるように



してもよい。

【0010】またさらに、前記検索過程は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索するようにしてもよい。また、前記特許文献は、複数の「請求項」の記載を含み、前記検索過程は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析過程を備えるようにしてもよい。さらに前記請求項解析過程は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出するようにしてもよい。

【0011】さらにまた、前記請求項解析過程は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別するようにしてもよい。また、前記特許文献は、「請求項」の記載を含み、前記検索過程は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析過程を備えるようにしてもよい。

【0012】また、検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有するコンピュータを特許検索装置として機能させる制御プログラムは、検索元特許文献を解析させ、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出させ、前記検索用語句および前記検索条件語句を比較させ、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索させる、ことを特徴としている。この場合において、各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出させるようにしてもよい。また、前記特許文献は、複数の「請求項」の記載を含み、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出させるようにしてもよい。さらに前記特許文献は、「請求項」の記載を含み、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出させるようにしてもよい。また、上記各制御プログラムを記録媒体に記録するようにしてもよい。

【0013】

【発明の実施の形態】次に本発明の好適な実施の形態について図面を参照して説明する。本実施形態は、本発明を特許検索システムに適用する場合のものである。

〔1〕特許検索データベース登録システム

まず特許検索システムに用いられる特許検索データベースを構築するための特許検索データベース登録システムについて説明する。

〔1.1〕特許検索データベース登録システムの構成

図1は特許検索データベース登録システムの概要構成ブロック図である。特許検索データベース登録システム10は、大別すると、各種データをデータベースとして蓄積するデータベース部11と、データベース部11に各種データを登録するための登録処理部12と、を備えている。ここで、特許検索データベース登録システム10は、コンピュータシステムにおいて実現可能であり、登録処理部12の機能は、登録処理部12を構成する各部に対応する、マイクロプロセッサで実行可能なプログラムによって実現される。また、このようなプログラムは、半導体メモリ、CD-ROMなどの記録媒体から直接実行してもよい。また、外部記憶装置に予めプログラムインストールして実行することも可能である。さらにプログラムの実行に先立って実行する毎、あるいは、最初に一度だけ、インターネットなどのネットワークを介してインストールするようにしてもよい。

【0014】〔1.1.1〕データベース部の構成

まず、データベース部の構成について説明する。データベース部11は、ハードディスクなどの外部記憶装置に構築されている。そしてデータベース部11は、大別すると、単語データベース15と、構成要件（発明特定事項）データベース16と、従属関係データベース17と、参照文献データベース18と、を備えている。単語データベース15は、特許文献PDに含まれている単語あるいは複合語を予め指定された特定の分類（分野）に関連させて記憶させるためのデータベースである。

【0015】図2は単語データベースの基本構成である。単語データベース15は、大別すると、ページテーブル21と、キーワードテーブル22と、ワードテーブル23と、を備えている。図2においては、一つのページテーブル21に対し、1系統のキーワードテーブル22およびワードテーブル23のみを図示しているが、実際には、一つのページテーブル21に対し、複数系統のキーワードテーブル22およびワードテーブル23が構成され、全体として、ツリー構造（木構造）をなしている。ページテーブル21は、ページIDデータ31、認識番号データ32、タイトルデータ33、更新日時データ34、書誌事項データ35および被リンク数データ36を備えている。なお、このデータ構成は一例であり、これに限られるものではない。例えば、対応する公報のページ数や技術分野などのデータを組み込むようにしても構わない。

【0016】ページIDデータ31は、登録する特許文献（特許文書）PDのそれぞれに固有（unique）に割り当てられ、各特許文献PDを特定するためのデータである。認識番号データ32は、公報番号など特許庁によって各特許文献PDに割り振られた値を有するデータである。タイトルデータ33は、各特許文献PDの内容を表す名称を格納している。例えば、特許文献PDが特許公報である場合には、発明の名称である。更新日時データ

34は、各特許文献PDの登録日時を格納している。書誌事項データ35は、特許文献PDが特許公報である場合には、公報発行国、公報種別、公開番号、公開日、国際特許分類情報（FI分類番号等）、審査請求の有無、請求項の数、出願形態、全頁数、出願番号、出願日、優先権主張基礎出願の出願番号、優先権主張基礎出願の出願番号、優先権主張基礎出願の出願日、出願人情報（識別番号、氏名または名称、住所または居所）、発明者情報（氏名、住所または居所）、代理人情報（識別番号、氏名、資格）、テーマコード、Fターム等である。すなわち、書誌事項データ35は、各特許文献PDに関わる様々な書誌的な情報を格納するためのデータである。被リンク数データ36は、各特許文献PDが参照されている他の特許文献の文献数を格納するためのデータである。

【0017】また、キーワードテーブル22は、ページテーブル21とワードテーブル23とを関係付けするためのデータである。そして、キーワードテーブル22は、ページIDデータ41、ワードIDデータ42、出現回数（cost）データ43、重要度データ44および段落フラグデータ45を備えている。ページIDデータ41は、ページテーブル21のページIDデータ31に対応するデータが格納するデータである。すなわち、ページテーブル21のページIDデータ31と同一のデータが格納されていれば、当該キーワードテーブル22は、同一のページIDデータを有するページテーブル21に関連するデータを格納していることを表している。ワードIDデータ42は、対応する単語についての情報を格納しているワードテーブル23との関連づけを行うためのデータである。

【0018】出現回数（cost）データ43は、ワードIDデータ42で特定される単語のページIDデータ41で特定される特許文献中における出現回数に相当するデータである。すなわち、ある単語が当該特許文献中に何回使用されたかを表すデータである。重要度データ44は、ページIDデータ41で特定される特許文献におけるワードIDデータ42で特定される単語の重要度としてTFIDF法により算出したTFIDF値を格納している。このTFIDF値は、後述するワードテーブルに格納されているIDFデータ（IDF値；全特許文献中における対応する単語の出現した文献数に相当）および出現回数データ43に基づいて算出される。すなわち、ある単語が出現した文献数が少ないほど、かつ、各文献における当該単語の出現数が多いほど重要な単語であると認識されることとなる。

【0019】段落フラグデータ45は、ワードIDデータ42で特定される単語がページIDデータ41で特定される特許文献中のいずれの段落で検出されたかを示すデータである。この場合における段落とは、特許文献が論文や書籍などの通常の文献の場合には、一般的な段落

の意味の他、章や節などの概念を含むものである。また、特許文献が特許公報などの場合には、発明の名称の欄、特許請求の範囲の欄、発明の詳細な説明の欄、図面の簡単な説明の欄、要約の欄などについても段落として扱っている。

【0020】ワードテーブル23は、単語データ51、ワードIDデータ52およびIDFデータ53を備えている。単語データ51は、形態素解析後の単語が登録されている。ワードIDデータ52は、単語データ51として登録されている単語を特定するためのデータである。IDFデータ53は、単語データ51として登録されている単語が全特許文献中において出現した文献数に相当するIDF値が登録されている。構成要件データベース16は、特許文献である特許公開公報、特許掲載公報などから、特許請求の範囲に含まれる構成要件（いわゆる発明特定事項）のデータを記憶している。具体的には、各構成要件の名称、各構成要件の作用を特徴づけるような単語などを格納している。従属関係データベース17は、請求項（クレーム）間の繋がりである従属関係に関するデータを記憶している。

【0021】図3に従属関係データベース17のデータフォーマットを示す。従属関係データベース17は、ページIDデータ55と、被引用請求項データ56と、引用請求項データ57と、を備えている。ページIDデータ55は、ページテーブル21のページIDデータ31と同様のデータであり、ページIDデータ55に対応する被引用請求項データ56および引用請求項データ57は、このページIDデータ55に対応する特許文献に含まれる請求項の従属関係を表すこととなる。被引用請求項データ56は、他の請求項により引用される請求項の請求項番号が登録される。引用請求項データ57は、他の請求項を引用する請求項の請求項番号が登録される。

【0022】具体的には、ページデータ=1で特定される特許文献の特許請求の範囲の記載において、請求項1および請求項2が独立項となっており、請求項3が請求項1および請求項2を引用した従属項となっており、ページデータ=2で特定される特許文献の特許請求の範囲の記載において、請求項1が独立項となっており、請求項2が請求項1を引用した従属項となっている場合には、従属関係データベース17には、図4に示すようなデータが登録される。すなわち、図4に示すように、以下の①～③のデータが登録される。

①ページIDデータ=1、被引用請求項データ=1、引用請求項データ=3

②ページIDデータ=1、被引用請求項データ=2、引用請求項データ=3

③ページIDデータ=2、被引用請求項データ=1、引用請求項データ=2

【0023】参照文献データベース18は、各特許文献について、当該文献中で参照している参照文献あるいは

審査などにおいて参照された参照文献に関するデータを格納している。具体的には、参照文献番号（特許公開番号、特許番号、……）、文献名、文献参照頁、文献発行元、著者名などが必要に応じて格納されることとなる。参照文献データベース18は、図8に示すように、参照先ページIDデータ81、参照元ページIDデータ82、タイトルデータ83、文献種別データ84および文献書誌情報データ85を備えている。参照先ページIDデータ81は、参照先の特許文献を特定するためのページIDデータである。参照元ページIDデータ82は、参照元の特許文献を特定するためのページIDデータである。

【0024】タイトルデータ83は、参照先の文献タイトルを登録する。上述の例の場合には「特開平△-12345号公報」が文献タイトルとなる。文献種別データ84は、当該参照先の特許文献の種別を表すデータである。例えば、参照先の特許文献の種別が国内特許公報の場合には文献種別データ=0、外国特許公報の場合には文献種別データ=1、論文である場合には文献種別データ=2のように設定される。文献書誌情報データ85は、文献名、文献参照頁、文献発行元、著者名などの書誌情報が所定の順番でCSV形式などのテキストデータ形式で格納されている。なお、この文献書誌情報データ85は、これらの情報が必要であれば、必ずしも設ける必要はない。

#### 【0025】[1.1.2] 検索処理部の構成

次に検索処理部12の構成について再び図1を参照して説明する。検索処理部12は、大別すると、形態素解析部61と、重要度算出部62と、請求項解析部63と、形態素解析用辞書部64と、参照文献抽出部65と、標準化部66と、ストップワード処理部67と、を備えている。形態素解析部61は、特許文献PDを形態素解析して頻出する名詞、サ変名詞、その他の名詞あるいはサ変名詞に類する単語を単語データベース15に登録する。この形態素解析部61による単語の単語データベース15への登録に先立って、標準化部66は、単語の標準化（統一化）を行う。例えば、単語として「PC」、「パソコン」、「パーソナルコンピュータ」の3種類の単語が形態素解析により得られた場合に、これらの単語は同一の意味を有しているので「パソコン」の単語に統一するのである。また、ストップワード処理部67は、特許分野において特許文献の種類に関わらず頻出する単語の単語データベース15への登録を行わないように処理をしている。例えば、単語として「請求項」、「出願」、「特許」、「本発明」、「実用新案」等の単語については、単語データベース15への登録を行わないようにしている。

【0026】重要度算出部62は、形態素解析されて登録された単語に対して重要度をTFIDF法を用いてTFIDF値として計算し、単語データベース15へ登録

する。請求項解析部63は、大別すると、従属関係解析部63Aと、構成要件抽出部63Bと、を備えている。従属関係解析部63Aは、特許文献が特許公報や実用新案公報などである場合に、特許請求の範囲や実用新案登録請求の範囲などを解析し、各請求項同士の従属関係を調べる。また、構成要件抽出部63Bは、特許請求の範囲や実用新案登録請求の範囲などから構成要件（例えば、発明特定事項）を抽出する。形態素解析用辞書部64は、形態素解析部61が形態素解析を行う際に用いる一般的に知られているような態様の形態素解析辞書を格納している。参照文献抽出部65は、各特許文献について、当該文献中で参照している参照文献あるいは審査などにおいて参照された参照文献を抽出し、参照文献データベース18に登録する。

#### 【0027】[1.2] 特許検索データベース登録システムの動作

次に特許検索データベース登録システムの動作を説明する。以下の説明においては、特許文献として特許公報を登録対象とする場合を例として説明する。図5に特許検索データベース登録処理のフローチャートを示す。まず特許検索データベース登録システム10は、形態素解析を行い（ステップS1）、単語データベースに登録する（ステップS2）。具体的には、形態素解析部61は、特許文献を形態素解析するに際し、各特許文献毎に固有の番号を所定の条件に基づいてページIDとして割り振り、ページテーブルのページIDデータおよびキーワードテーブルのページIDデータに登録する。次に形態素解析部61は、特許文献である特許公報の形態素解析を行い、公報番号を認識番号データ32として登録する。

【0028】続いて、形態素解析部61は、発明の名称を抽出し、当該発明の名称をタイトルデータ33として登録する。そして形態素解析部61は、当該特許公報について、当該データベース登録システムにおける登録日時を更新日時データに登録する。また、形態素解析部61は、特許公報から、公報発行国、公報種別、公開番号、公開日、国際特許分類情報、審査請求の有無、請求項の数、出願形態、全頁数、出願番号、出願日、優先権主張基礎出願の出願番号、優先権主張基礎出願の出願番号、優先権主張基礎出願の出願日、出願人情報、発明者情報、代理人情報、テーマコード、Fタームなどを抽出し、書誌事項データ35として登録する。次に特許検索データベース登録システムは、形態素解析の結果に基づいて各単語の重要度を算出し（ステップS3）、単語データベース15のキーワードテーブル22に登録する（ステップS2）。

【0029】この場合において、重要度算出部62は、キーワードテーブルの出現回数（cost）データ43およびワードテーブルのIDFデータ53に基づいて、重要度データとして格納する重要度を算出している。具体的には、登録対象の全特許文献に含まれる単語（キーワー

10

20

30

40

50

ド) 毎の出現回数をTFとし、登録対象の全特許文献数をNとし、登録対象の全特許文献のうち重要度算出対象の単語が含まれる特許文献数をnとした場合に、次式により重要度データとしてTFIDF値を算出する。

$$IDF = \log(N/n)$$

$$TFIDF = TF \cdot IDF$$

この場合において、算出したIDFは、IDFデータとしてワードテーブルに登録される。次に特許検索データベース登録システムは、請求項の解析の一部である請求項間の従属関係を解析し(ステップS4)、従属関係データ

ベース17に登録する(ステップS5)。  
【0030】この場合において、従属関係解析部63Aは、特許請求の範囲や実用新案登録請求の範囲における「請求項〇〇記載の□□方法において」、「……ことを特徴とする請求項△△記載の■装置。」などの記載に基づいて、各請求項同士の従属関係を解析し、図4に示したような従属関係を従属関係データベースに登録する。次に特許検索データベース登録システムは、請求項の解析の一部である構成要件抽出を行う(ステップS6)。この場合において、構成要件抽出部63Bは、請求項の記載から構成要件(発明特定事項)を抽出し、構成要件データベース16に登録する(ステップS7)。具体的に、以下の請求項の記載から構成要件を抽出する場合について説明する。

【0031】[請求項1] 時系列に順序づけられたn個の文書を関連づける方法であって、前記n個の文書間の類似度を計算するステップと、前記類似度から時間制約を用いて類似度行列を作成するステップと、前記類似度行列を前記文書の関連づけを示す隣接行列に変換するステップと、を有することを特徴とする文書の関連づけ方法。

【0032】この場合には、「ステップ」というキーワードにより以下の3つの構成要件が抽出される。

「前記n個の文書間の類似度を計算するステップ」

「前記類似度から時間制約を用いて類似度行列を作成するステップ」、

「前記類似度行列を前記文書の関連づけを示す隣接行列に変換するステップ」

そして抽出された構成要件は、構成要件データベース16に登録されることとなる。

【0033】具体的には、構成要件データベース16は、図6に示すように、ページIDデータ71、クレームIDデータ72および構成要件テキストデータ73を備えている。ページIDデータ71は、ページテーブル21のページIDデータ31と同様のデータであり、ページIDデータ71に対応するクレームIDデータ72および構成要件テキストデータ73は、このページIDデータ71に対応する特許文献に含まれる請求項の内容を表すこととなる。クレームIDデータ72は、構成要件テキストデータ73に格納されているテキストデータ

のクレーム番号を特定するためのデータである。

【0034】構成要件テキストデータ73は、構成要件をテキストデータとして格納するためのデータである。そして、上記請求項を構成要件データベース16に登録した場合には、図7に示すようになる。すなわち、特許文献を特定するための番号=231がページIDデータ71として格納され、請求項1を表す番号=1がクレームIDデータ72として格納され、請求項1の内容に相当するテキストデータ=「前記n個の文書間の類似度を計算するステップ」が構成要件テキストデータ73として格納される。

【0035】次に特許検索データベース登録システム10は、参照文献を抽出する(ステップS8)。この場合において参照文献抽出部65は、特許公報(特許文献)内において、いずれかの特許文献(国内外の特許公報、論文等)を引用しているかを検出し、引用先の情報(公報番号、書名、論文タイトルなど)を参照先データベースに登録する。また、参照先の特許文献が既に登録されている場合には、当該参照先の特許文献に対応するページテーブル21の被リンク数データ36の値を更新する。

【0036】具体的には、特許文献である特許公報中に「特開平△-12345号公報によると……」や「文献“Fast Algorithms for Mining Association”を参照……」などの記載を抽出し、参照文献データベース18に登録する。

【0037】[1.3] 特許検索システムの構成

図9は特許検索システムの概要構成ブロック図である。

図9において、図1の特許検索データベース登録システムと同様の部分には同一の符号を付し、その詳細な説明を省略する。特許検索システム100は、大別すると、各種データをデータベースとして蓄積するデータベース部11と、データベース部11を用いて特許検索を行う検索処理部102と、ディスプレイ部103と、入力部104と、を備えている。

【0038】ここで、特許検索システム100は、コンピュータシステムにおいて実現可能であり、検索処理部102の機能は、検索処理部102を構成する各部に対応する、マイクロプロセッサで実行可能なプログラムによって実現される。また、このようなプログラムは、半導体メモリ、CD-ROMなどの記録媒体から直接実行してもよい。また、外部記憶装置に予めプログラムインストールして実行することも可能である。さらにプログラムの実行に先立って実行する毎、あるいは、最初に一度だけ、インターネットなどのネットワークを介してインストールするようにしてもよい。検索処理部102は、入力部104を介して入力された検索元特許文献(あるいは検索元特許文献を特定するための特定情報)に基づいて後述する類似検索、クレーム検索あるいは参照先検索を行う。ディスプレイ部103は、検索結果を

含む各種データを表示し、検索オペレータに対して提示する。入力部104は、キーボード、マウス、タブレット、スキャナ、リムーバブルディスク装置、通信インタフェース部などで構成され、検索元特許文献に対応するデータの入力や各種検索条件などの設定を行う。

【0039】[1. 3. 1] 検索処理部の構成  
次に検索処理部の構成について説明する。検索処理部102は、大別すると、構文解析・形態素解析部111と、形態素解析用辞書部112と、図表作成部113と、実検索処理部114と、標準化部115と、ストップワード処理部116と、を備えている。構文解析・形態素解析部111は、語句抽出部として機能し、形態素解析部121と、請求項解析部122と、参照文献抽出部123と、を備えている。形態素解析部121は、形態素解析用辞書64を用いて、特許文献を形態素解析して頻出する名詞、サ変名詞、その他の名詞あるいはサ変名詞に類する単語を抽出する。

【0040】この形態素解析部121による形態素解析結果の出力に際して、標準化部115は、単語の標準化（統一化）を行う。また、ストップワード処理部116は、特許分野において特許文献の種類に関わらず頻出する単語について形態素解析結果として出力しないように処理をおこなっている。請求項解析部122は、検索元特許文献が特許公報や実用新案公報などである場合に、特許請求の範囲や実用新案登録請求の範囲などを解析し、各請求項同士の従属関係を調べる。また、特許請求の範囲や実用新案登録請求の範囲などから構成要件（例えば、発明特定事項）を抽出する。参照先抽出部123は、検索元特許文献について、当該検索元特許文献中で参照している参照文献あるいは審査などにおいて参照された参照文献を抽出する。形態素解析用辞書部112は、形態素解析部121が形態素解析を行う際に用いる形態素解析辞書を格納している。

【0041】図表作成部113は、検索結果を図表化して検索オペレータに対して提示する。実検索処理部114は、検索部として機能し、大別すると、類似検索部131と、クレーム検索部132と、参照先検索部133と、を備えている。類似検索部131は、検索元特許文献と類似の特許文献を検索し、類似関係にある特許文献を抽出し、その類似関係を把握する。クレーム検索部132は、検索元特許文献のクレーム（請求項）の従属関係を把握する。参照文献検索部133は、検索元特許文献において参照している特許文献あるいは審査などにおいて参照された特許文献を抽出する。

【0042】[1. 4] 特許検索システムの動作  
次に特許検索システムの動作を説明する。以下の説明においては、検索元特許文献として特許公報を用いる場合を主として説明する。図10に特許検索処理のフローチャートを示す。この場合において、特許検索処理に先立って、オペレータにより類似の特許文献を検索する類似

検索処理を行うか、あるいは、当該特許文献の請求項相互間の関係を調べるクレーム検索処理を行うかのいずれかが指定されているものとし、類似検索処理を行う場合には類似検索フラグがセットされ、クレーム検索処理を行う場合には請求項フラグがセットされているものとする。まず特許検索システム100は、検索元特許文献がデータベース部11に未登録否かを判別する（ステップS11）。

【0043】ステップS11の判別において、検索元特許文献が既にデータベースに登録されている場合には（ステップS11；No）、類似検索フラグあるいは請求項フラグのいずれかがセットされているかを判別する（ステップS12）。ステップS12の判別において、類似検索フラグがセットされている場合には（ステップS12；類似フラグセット）、実検索処理部114の参照先検索部133は、参照文献データベース18を検索して検索元特許文献が参照している特許文献を抽出する（ステップS13）。次に実検索処理部114の類似検索部131は、単語データベース15を参照して検索元特許文献に含まれる単語を読み出し、各単語毎に類似検索を行い（ステップS14）、検索元特許文献と単語データベース15に登録されている他の特許文献との間の距離計算および請求項の構成要件についてパターンマッチングを行う。

【0044】より具体的には、検索元特許文献が日本国内の特許公報である場合には、例えば、発明の詳細な説明および要約に含まれる単語のそれぞれをベクトル表現し、単語データベースに登録されている他の特許文献に含まれる単語に対応するベクトルとの間で距離計算処理を行う。また、検索元特許文献において特許請求の範囲を構成する請求項の構成要件と、構成要件データベースに登録されている構成要件との間のパターンマッチング処理を行い、類似度を求める。この距離計算処理およびパターンマッチング処理により得られる類似の度合いに基づいて所定の類似範囲内にある特許文献を特定する。また、同様の距離計算処理およびパターンマッチング処理を検索元特許文献が参照している特許文献についても行う。そして、類似検索処理が終了すると、図表作成部113は、類似検索結果を図表化するための処理を行う（ステップS15）。図表作成部113において、類似検索結果を図表化するための処理が終了するとディスプレイ部103において類似検索結果に対応する図表を表示する（ステップS16）。

【0045】図11に類似検索結果に対応する図表の一例を示す。なお、以下の説明において、特許公報という表現は、日本国における特許掲載公報ばかりでなく、公開特許公報その他の公報も含むものである。図11において、特許公報Xが検索元特許文献である。特許公報Xに対して直接あるいは間接に結ばれている特許文献は、直接あるいは間接的に参照されている特許文献である。



またいずれの特許文献にも結ばれていない特許文献は、類似文献であるが、直接的な参照関係にはない特許文献である。この場合には、特許文献間の 2 次元的距離が近いほど類似度が高いようになっている。さらに特許公報 X の表示位置に対して左側に表示され、かつ、特許公報 X に直接結ばれている特許文献（米国特許公報 S、特許公報 C、特許公報 B、特許公報 D、大学論文 R）は、特許公報 X が直接参照している特許文献である。すなわち、米国特許公報 S、特許公報 C、特許公報 B、特許公報 D、大学論文 R は、検索元特許文献である特許公報 X 中において直接的に記述され、参照されている特許文献である。

【0046】さらにこれらの米国特許公報 S、特許公報 C、特許公報 B、特許公報 D、大学論文 R と特許公報 X との間の 2 次元的距離が近いほど類似度が高いようになっている。また、各特許文献を表す枠内に表示されている数字は、当該特許文献を参照している特許文献数である。この図の例の場合には、特許公報 A および特許公報 E は、参照件数が多いので、基本特許の可能性が高いということが判断できる。従って、特許公報 X が特許公開公報であるような場合には、特許出願人は、特許公報 A および特許公報 E に記載の発明をはじめに参照すべきと考えられることとなる。さらに特許公報 X に類似している特許公報 N、Q、P も存在しているので、これらの特許公報に記載の発明についても注意が必要であることがわかる。この場合において、ディスプレイ部 103 の表示画面上において、いずれかの特許文献を選択し、検索種別を指定すれば、同様に当該選択した特許文献を検索元特許文献とする検索処理がなされる。このように構成することにより、階層的に検索を継続することができ、

所望の特許文献を効率的に検索することができる。

【0047】また、複数の特許文献を線分で結ぶ場合に線分の色分けを行って系統分けをしたり、出願人毎や、発明者毎、代理人（弁理士など）毎に特許文献を表す枠内の色を変更したり、ブリンク（点滅）表示させることにより、より特許文献相互間の関係を把握しやすくすることが可能となる。一方、ステップ S12 の判別において、請求項フラグがセットされている場合には（ステップ S12；請求項フラグセット）、実検索処理部 114 のクレーム検索部 132 は、構成要件データベースおよび従属関係データベースを検索して請求項相互の従属関係を抽出する（ステップ S17）。そして、クレーム検索処理が終了すると、図表作成部 113 は、クレーム検索結果を図表化するための処理を行う（ステップ S15）。図表作成部 113 において、クレーム検索結果を図表化するための処理が終了するとディスプレイ部 103 においてクレーム検索結果に対応する図表を表示する（ステップ S16）。

【0048】図 12 にクレーム検索結果に対応する図表の一例を示す。この場合において、各請求項には、以下

のような記載があるものとする。

請求項 1： 「……を備えたことを特徴とする知識抽出方法」

請求項 2： 「請求項 1 記載の知識抽出方法において、……」

請求項 3： 「請求項 2 記載の知識抽出方法において、……」

請求項 4： 「……を備えたことを特徴とする知識抽出装置」

請求項 5： 「請求項 4 記載の知識抽出装置において、……」

請求項 6： 「請求項 4 記載の知識抽出装置において、……」

請求項 7： 「請求項 4 記載の知識抽出装置において、……」

請求項 8： 「請求項 4 記載の知識抽出装置において、……」

請求項 9： 「……を備えたことを特徴とする知識抽出プログラム」

【0049】クレーム検索部 132 は、各請求項の記載に基づいて、請求項 1、請求項 4 および請求項 9 は独立請求項であることを検出する。すなわち、当該請求項の記述中に他の請求項の記述を含まない請求項あるいは発明の名称が異なる他の請求項の記述のみを含む請求項である場合に独立請求項であるとする。また、クレーム検索部 132 は、当該請求項より前に記述されている請求項の発明の名称と発明の名称が同一であり、かつ、他の請求項の記載（例えば、「請求項○○」）を含む請求項は、当該他の請求項に従属する従属請求項であると判断する。すなわち、請求項 2 は請求項 1 の従属請求項、請求項 3 は請求項 2 の従属項、請求項 5～請求項 8 はそれぞれ請求項 4 の従属項であることを検出する。この結果、図 12 に示すように、請求項 1 に対応する枠と請求項 2 に対応する枠は線分により接続され、さらに請求項 2 に対応する枠と請求項 3 に対応する枠は線分により接続されて表示される。また、請求項 4 に対応する枠には、請求項 5～請求項 8 に対応する枠が並列的に線分により接続されて表示される。さらに請求項 9 に対応する枠は独立して表示される。このような表示がなされることにより、検索者は、直感的に当該請求項群の中で基本的な内容を特定している独立請求項を把握できるとともに、請求項間の従属関係を容易に把握できる。

【0050】[1.5] 実施形態の効果

以上の説明のように、本実施形態によれば、特許検索を行うに際し、オペレータがキーワードを入力することなく、「漏れの少ない」調査を行い、かつ、不必要な情報が含まれることを抑制することができる。また、複数の特許文献間の参照、被参照関係を明確に提示し、さらに参照関係がなくても近似する技術内容の特許文献を抽出することができる。さらにまた、同一の特許文献中にお

ける請求項の従属関係を明確にし、オペレータに対して視覚的に容易に把握できる状態で提示できる。

# 【0051】[2]実施形態の変形例

## [2.1]第1変形例

以上の説明においては、データベース部11を登録処理部12（あるいは検索処理部102）と一体に構成していたが、両者をネットワークを介して分散処理システムとして構成することも可能である。この場合において、さらにデータベース部11を構成する各データベース15、16、17、18および形態素解析用辞書64をネットワークを介して別のデータベースサーバに格納するように構成し、複数の登録処理部12（あるいは検索処理部102）として機能するコンピュータシステムから利用可能な構成とすることも可能である。

## 【0052】[2.2]第2変形例

以上の説明においては、標準化部66（あるいは標準化部115）を必須の構成として説明したが、必ずしも標準化部23を設けなくてもデータベースの容量は多少増加するというデメリットはあるが、ほぼ同様な効果を得ることが可能である。

## 【0053】

【発明の効果】本発明によれば、特許検索を行うに際し、「漏れの少ない」検索を行え、かつ、不必要な情報が含まれることを抑制することができる。また、複数の特許文献間の参照、被参照関係を明確に提示し、さらに参照関係がなくても近似する技術内容の特許文献を抽出することができる。さらにまた、同一の特許文献中における請求項の従属関係を明確化し、提示することができる。

## 【図面の簡単な説明】

【図1】 特許検索データベース登録システムの概要構成ブロック図である。

【図2】 単語データベースの基本構成図である。

【図3】 従属関係データベースのデータフォーマットの説明図である。

【図4】 従属関係データベースのデータ登録例の説明図である。

【図5】 特許検索データベース登録処理のフローチャートである。

【図6】 構成要件データベースのデータフォーマットの説明図である。

\*【図7】 構成要件データベースのデータ登録例の説明図である。

【図8】 参照文献データベースのデータフォーマットの説明図である。

【図9】 特許検索システムの概要構成ブロック図である。

【図10】 特許検索処理のフローチャートである。

【図11】 類似検索結果に対応する図表の一例の説明図である。

10 【図12】 クレーム検索結果に対応する図表の一例の説明図である。

## 【符号の説明】

10……特許検索データベース登録システム

11……データベース部

12……登録処理部

15……単語データベース

16……構成要件（発明特定事項）データベース

17……従属関係データベース

18……参照文献データベース

20 21……ページテーブル

22……キーワードテーブル

23……ワードテーブル

31……ページIDデータ

32……認識番号データ

33……タイトルデータ

34……更新日時データ

35……書誌事項データ

36……被リンク数データ

30 100……特許検索システム

102……検索処理部

103……ディスプレイ部

104……入力部

111……構文解析・形態素解析部

112……形態素解析用辞書部

113……図表作成部

114……実検索処理部

115……標準化部

116……ストップワード処理部

121……形態素解析部

40 122……請求項解析部

\* 123……参照文献抽出部

【図3】

17	55
pageID	55
被引用請求項データ	56
引用請求項データ	57

【図4】

17	①	②	③	...
55	1	1	2	...
56	1	2	1	...
57	3	3	2	...

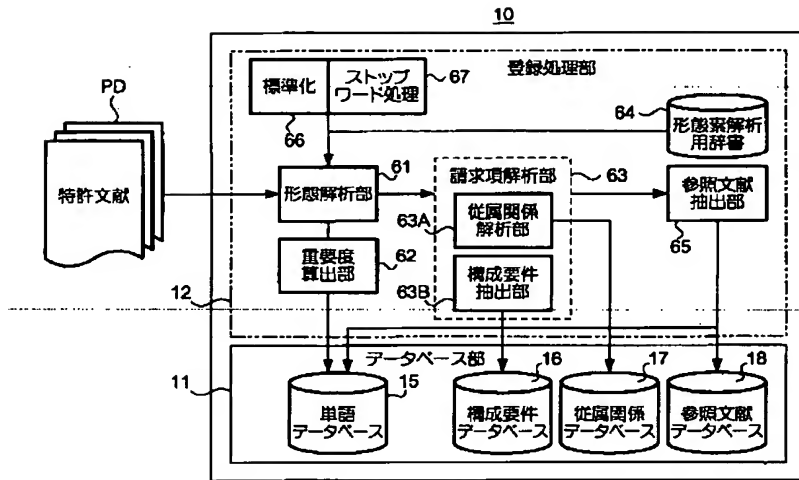
【図6】

16	71
pageID	71
クレームID	72
構成要件テキスト	73

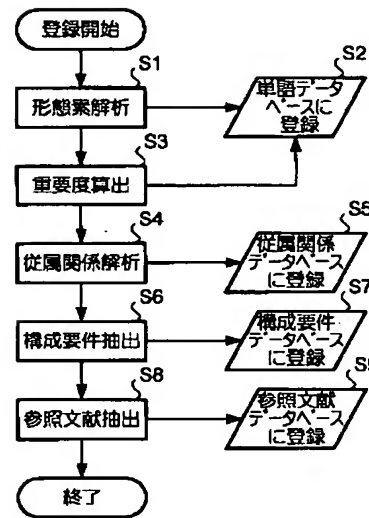
【図7】

71	231
72	1
73	前記n個の文書間の類似度を計算するステップ

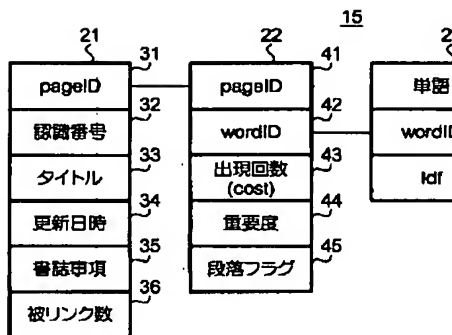
【図1】



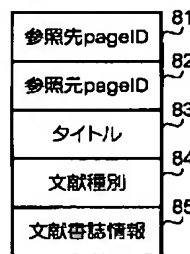
【図5】



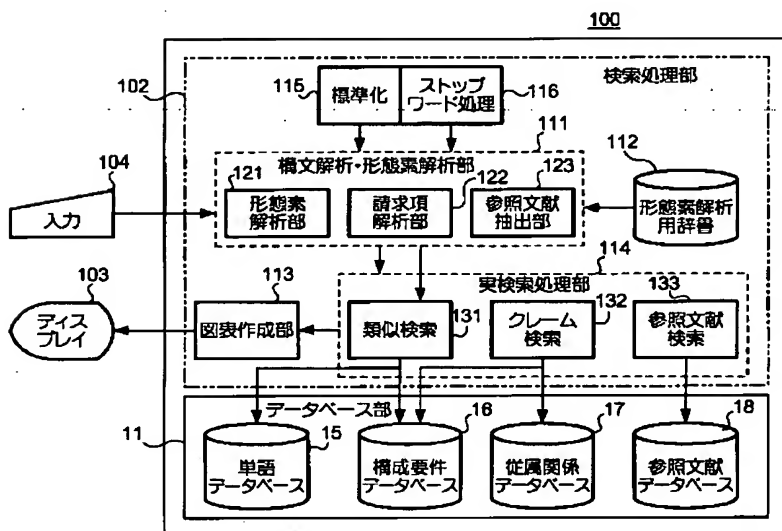
【図2】



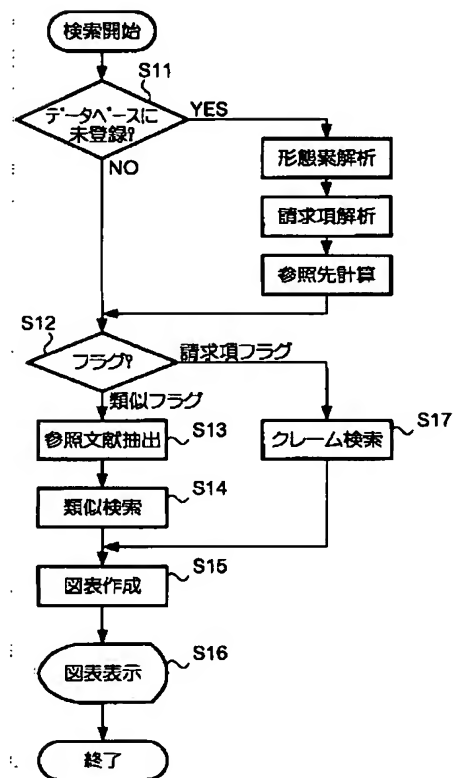
【図8】



【図9】

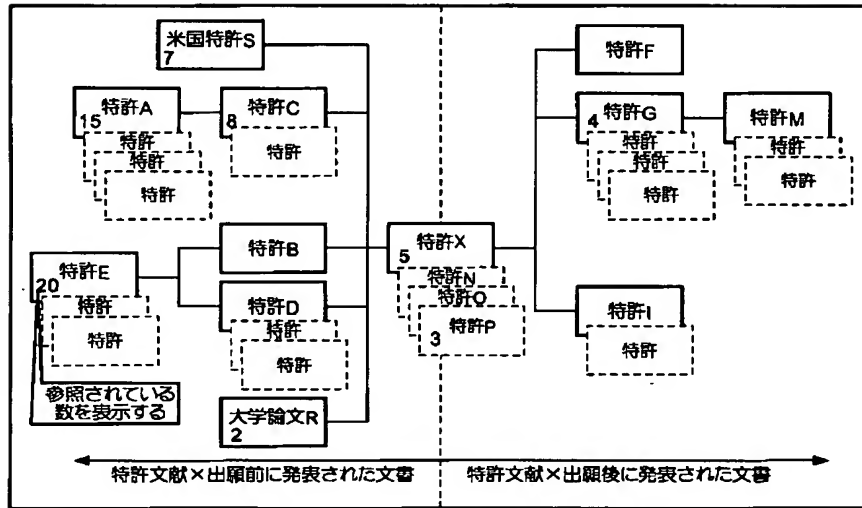


【図10】

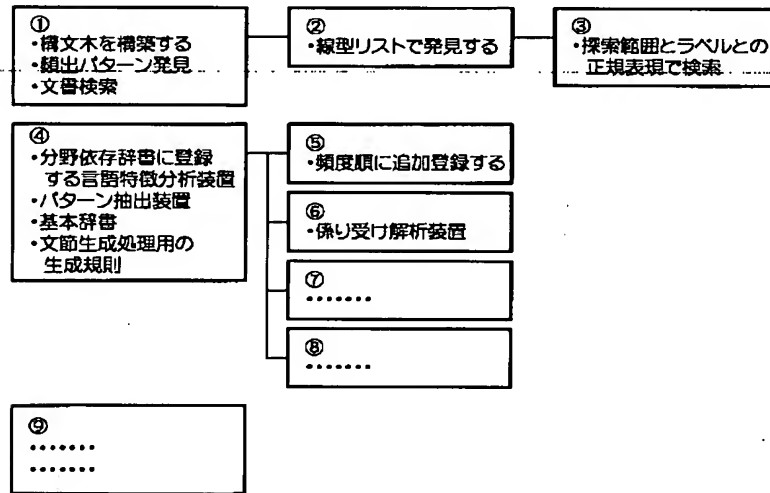




【図11】



【図12】



【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第3区分  
 【発行日】平成17年6月23日(2005.6.23)

【公開番号】特開2003-157262(P2003-157262A)  
 【公開日】平成15年5月30日(2003.5.30)  
 【出願番号】特願2001-355274(P2001-355274)  
 【国際特許分類第7版】

G 0 6 F 17/30

【F I】

G 0 6 F 17/30 1 7 0 A

G 0 6 F 17/30 1 4 0

G 0 6 F 17/30 2 1 0 A

G 0 6 F 17/30 3 5 0 C

G 0 6 F 17/30 3 6 0 Z

【手続補正書】

【提出日】平成16年9月28日(2004.9.28)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

【補正の内容】

【発明の名称】特許検索装置、特許検索装置の制御方法、制御プログラムおよび記録媒体、並びに検  
置、検索装置の制御方法、制御プログラム

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部と、

検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出部と、

前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索部と、  
 を備えたことを特徴とする特許検索装置。

【請求項2】

請求項1記載の特許検索装置において、

各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出部を備えたことを特徴とする特許検索装置。

【請求項3】

請求項2記載の特許検索装置において、

前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備えたことを特徴とする特許検索装置。

【請求項4】

請求項3記載の特許検索装置において、

前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成部と、  
前記図表作成部により作成された図表を提示するディスプレイ部と、  
を備えたことを特徴とする特許検索装置。

【請求項 5】

請求項 1 ないし請求項 4 のいずれかに記載の特許検索装置において、  
前記検索元特許文献の形態素解析を行って前記単語あるいは前記複合語を抽出する形態素解析部と、  
抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度算出部と、  
抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前記検索用語句として登録する登録部と、  
を備えたことを特徴とする特許検索装置。

【請求項 6】

請求項 1 ないし請求項 5 のいずれかに記載の特許検索装置において、  
前記検索部は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索することを特徴とする特許検索装置。

【請求項 7】

請求項 1 ないし請求項 6 記載の特許検索装置において、  
前記特許文献は、複数の「請求項」の記載を含み、  
前記検索部は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析部を備えたことを特徴とする特許検索装置。

【請求項 8】

請求項 7 記載の特許検索装置において、  
前記請求項解析部は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出することを特徴とする特許検索装置。

【請求項 9】

請求項 8 記載の特許検索装置において、  
前記請求項解析部は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別することを特徴とする特許検索装置。

【請求項 10】

請求項 1 ないし請求項 6 のいずれかに記載の特許検索装置において、  
前記特許文献は、「請求項」の記載を含み、  
前記検索部は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析部を備えたことを特徴とする特許検索装置。

【請求項 11】

検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有する特許検索装置の制御方法において、

検索元特許文献を解析し、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出過程と、

前記検索用語句および前記検索条件語句を比較し、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索する検索過程と

を備えたことを特徴とする特許検索装置の制御方法。

【請求項 12】

請求項 11 記載の特許検索装置の制御方法において、

各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出する参照先抽出過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項 13】

請求項 12 記載の特許検索装置の制御方法において、

前記特許検索データベース部は、抽出した前記参照関係に基づいて、前記検索元特許文書に対応づけて参照関係のある特許文献についての情報を格納する参照先データベース部を備え、

前記参照先データベース部の情報に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を表す図表を作成する図表作成過程と、

前記図表作成部により作成された図表を提示するディスプレイ過程と、

を備えたことを特徴とする特許検索装置の制御方法。

【請求項 14】

請求項 11 ないし請求項 13 のいずれかに記載の特許検索装置の制御方法において、

前記検索元特許文献の形態素解析を行って前記単語あるいは前記複合語を抽出する形態素解析過程と、

抽出された前記単語あるいは前記複合語の当該検索元特許文献における重要度を計算する重要度計算過程と、

抽出された前記単語あるいは前記複合語を前記重要度に対応づけて前記特許検索データベース部に前記検索用語句として登録する登録過程と、

を備えたことを特徴とする特許検索装置の制御方法。

【請求項 15】

請求項 11 ないし請求項 14 のいずれかに記載の特許検索装置の制御方法において、

前記検索過程は、前記検索用語句に対応するベクトルおよび前記検索条件語句に対応するベクトルの距離に基づいて関連性の高い特許文献を検索することを特徴とする特許検索装置の制御方法。

【請求項 16】

請求項 11 ないし請求項 15 記載の特許検索装置の制御方法において、

前記特許文献は、複数の「請求項」の記載を含み、

前記検索過程は、前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出する請求項解析過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項 17】

請求項 16 記載の特許検索装置の制御方法において、

前記請求項解析過程は、検索対象の「請求項」中に含まれる他の「請求項」の記載並びに当該検索対象の「請求項」および他の「請求項」の発明の名称に基づいて前記従属関係を抽出することを特徴とする特許検索装置の制御方法。

【請求項 18】

請求項 17 記載の特許検索装置の制御方法において、

前記請求項解析過程は、前記検索対象の「請求項」に前記他の「請求項」の記載があり、かつ、当該検索対象の「請求項」の発明の名称と前記他の「請求項」の発明の名称が同一である場合に、前記従属関係があると判別することを特徴とする特許検索装置の制御方法。

【請求項 19】

請求項 11 ないし請求項 15 のいずれかに記載の特許検索装置の制御方法において、

前記特許文献は、「請求項」の記載を含み、

前記検索過程は、前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出する請求項解析過程を備えたことを特徴とする特許検索装置の制御方法。

【請求項 20】

検索対象とする特許文献に含まれる単語あるいは複合語を予め検索用語句として当該特許文献に関連づけて記憶する特許検索データベース部を有するコンピュータを特許検索装

置として機能させる制御プログラムであって、

検索元特許文献を解析させ、当該条件設定用特許文献に含まれる単語あるいは複合語を検索条件語句として抽出させ、

前記検索用語句および前記検索条件語句を比較させ、前記検索元特許文献に対して類似性の高い特許文献を前記検索元特許文献に関連性の高い特許文献として検索させる、

ことを特徴とする制御プログラム。

【請求項 2 1】

請求項 2 0 記載の制御プログラムにおいて、

各特許文献中の記載に基づいて、前記検索元特許文献と前記関連性の高い特許文献との間の参照関係を抽出させる、

ことを特徴とする制御プログラム。

【請求項 2 2】

請求項 2 0 または請求項 2 1 記載の制御プログラムにおいて、

前記特許文献は、複数の「請求項」の記載を含み、

前記検索用語句に基づいて前記複数の「請求項」の相互間の従属関係を抽出させる、ことを特徴とする制御プログラム。

【請求項 2 3】

請求項 2 0 または請求項 2 1 記載の制御プログラムにおいて、

前記特許文献は、「請求項」の記載を含み、

前記検索用語句に基づいて前記「請求項」の記載に含まれる構成要件を抽出させる、ことを特徴とする制御プログラム。

【請求項 2 4】

請求項 2 0 ないし請求項 2 3 のいずれかに記載の制御プログラムを記録したことを特徴とする記録媒体。

【請求項 2 5】

検索対象とする文献に含まれる単語あるいは複合語を予め検索用語句として当該文献に関連づけて記憶する検索データベース部と、

検索元文献を解析し、当該条件設定用文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出部と、

前記検索用語句および前記検索条件語句を比較し、前記検索元文献に対して類似性の高い文献を前記検索元文献に関連性の高い文献として検索する検索部と、

を備えたことを特徴とする検索装置。

【請求項 2 6】

検索対象とする文献に含まれる単語あるいは複合語を予め検索用語句として当該文献に関連づけて記憶する検索データベース部を有する検索装置の制御方法において、

検索元文献を解析し、当該条件設定用文献に含まれる単語あるいは複合語を検索条件語句として抽出する語句抽出過程と、

前記検索用語句および前記検索条件語句を比較し、前記検索元文献に対して類似性の高い文献を前記検索元文献に関連性の高い文献として検索する検索過程と、

を備えたことを特徴とする検索装置の制御方法。

【請求項 2 7】

検索対象とする文献に含まれる単語あるいは複合語を予め検索用語句として当該文献に関連づけて記憶する検索データベース部を有するコンピュータを検索装置として機能させる制御プログラムであって、

検索元文献を解析させ、当該条件設定用文献に含まれる単語あるいは複合語を検索条件語句として抽出させ、

前記検索用語句および前記検索条件語句を比較させ、前記検索元文献に対して類似性の高い文献を前記検索元文献に関連性の高い文献として検索させる、

ことを特徴とする制御プログラム。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0052

【補正方法】変更

【補正の内容】

【0052】

[2.2] 第2変形例

以上の説明においては、標準化部66（あるいは標準化部115）を必須の構成として説明したが、必ずしも標準化部66を設けなくてもデータベースの容量は多少増加するというデメリットはあるが、ほぼ同様な効果を得ることが可能である。